

A Model for Machine Fault Diagnosis based on Mutual Exclusion Theory and Out-of-Distribution Detection

Peng Cui¹, Xuan Luo^{2*}, and Jing Liu¹

¹ School of Information Science and Engineer, Yanshan University, Qinhuangdao 066004, China
[e-mail: cptt0000@163.com, liujingjingr@163.com]

² Industrial Technology Center, Hebei Petroleum University of Technology, Chengde 067000, China
[e-mail: 81662638@qq.com]

*Corresponding author: Xuan Luo

*Received March 29, 2022; revised July 2, 2022; accepted August 25, 2022;
published September 30, 2022*

Abstract

The primary task of machine fault diagnosis is to judge whether the current state is normal or damaged, so it is a typical binary classification problem with mutual exclusion. Mutually exclusive events and out-of-domain detection have one thing in common: there are two types of data and no intersection. We proposed a fusion model method to improve the accuracy of machine fault diagnosis, which is based on the mutual exclusivity of events and the commonality of out-of-distribution detection, and finally generalized to all binary classification problems. It is reported that the performance of a convolutional neural network (CNN) will decrease as the recognition type increases, so the variational auto-encoder (VAE) is used as the primary model. Two VAE models are used to train the machine's normal and fault sound data. Two reconstruction probabilities will be obtained during the test. The smaller value is transformed into a correction value of another value according to the mutually exclusive characteristics. Finally, the classification result is obtained according to the fusion algorithm. Filtering normal data features from fault data features is proposed, which shields the interference and makes the fault features more prominent. We confirm that good performance improvements have been achieved in the machine fault detection data set, and the results are better than most mainstream models.

Keywords: out-of-distribution; convolutional neural network; mutually exclusive events; fusion networks; autoencoder.

1. Introduction

The primary goal of mechanical fault diagnosis is to determine whether or not the equipment is normal, which is a binary classification problem. One of the challenges of binary classification is distinguishing data near the decision boundary. To solve this problem, we consider the data to be two independent events A and B. The intersection of events A and B is empty. A and B are mutually exclusive events, known as mutually incompatible events [1], such as right and wrong, damaged and intact. Some classical single model neural networks, such as AlexNet [2], VGG [3], and ResNet [4], can identify each independent event. It prevents the emergence of decision boundaries and can validate each other to improve the model's overall performance. In recent years, among the models that have achieved good results in major competitions, multi-model fusion networks account for an increasing proportion. Multi-model fusion can fuse excellent models using scientific methods, overcome the bottleneck of a single model's generalization ability to unknown problems, and integrate the advantages of each model to obtain the optimal solution to the same problem [5]. This idea is worthy of being used for reference in mechanical fault diagnosis. All models can predict the input after training, but in many cases, they cannot judge whether the prediction results are accurate, which is unacceptable in practical application. A task directly linked to this problem is the out-of-distribution (OOD) detection task. OOD detection can identify In Distribution (ID) samples and OOD samples by the pre-trained model [6]. Therefore, the best way to improve machine fault diagnosis is to use OOD detection and model fusion based on mutual exclusion theory.

In this paper, we propose a homologous double model (HDM) method based on OOD detection and mutual exclusion principles. This method effectively improves the recognition effect of a single model and has a simple structure. We discarded the traditional process of extracting all types of features using the model, made the model focus on learning a kind of feature, and realized "let the professionals do the professional things." Autoencoders are often used for OOD detection, and a variational auto-encoder (VAE) is a randomly generated autoencoder that can provide calibration probability. It is also one of the few models that only learn one type of data. Therefore, when using VAE for binary classification, only half of the data can be used, which is bound to affect the performance of the model. We use VAE to train the data of two mutually exclusive events in the input data respectively, to obtain two pre-trained models. When detecting new input data, two pre-trained models are input, respectively, which can identify whether they belong to ID data. The smaller value is transformed into a correction value of another value according to the mutually exclusive characteristics, and finally, the classification result is obtained according to the fusion algorithm.

Our main contributions to this paper are as follows:

(1) A new HDM model is proposed based on the principle of extraterritorial detection and mutual exclusion, composed of two VAEs. The reconstruction rate of the two VAEs is obtained from the same data, and the final result is obtained after weighted calculation by the discriminator.

(2) When extracting low latitude features from normal data and fault data, a feature filtering method of fault data is proposed. Firstly, the normal data features are extracted. Then the fault data are mapped to low latitude to remove the features similar to the normal data features, shield the normal features, and highlight the fault features.

(3) The results of several experiments show that HDM effectively improves the recognition effect of a single model, and the recognition effect is better than the existing

classical network.

The rest of this paper is organized as follows: The related work is introduced in Section 2. The HDM model structure is presented in Section 3. Section 4 presents and analyzes the experimental results. The research is summarized in Section 5.

2. Related Work

2.1 Machine Fault Diagnosis

As a branch of machine learning, deep learning can automatically extract features from a large number of data, meet the requirements of adaptive feature extraction for mechanical fault diagnosis, effectively overcome the shortcomings of poor generalization ability and poor robustness of traditional manual feature extraction, and reduce the uncertainty of conventional fault diagnosis methods in the process of manual design and extraction. Tang et al. [7] summarized and analyzed the application of the rotating machinery fault diagnosis method based on a convolutional neural network. Haidong et al. [8] presented a novel method using a novel stacked transfer auto-encoder optimized by particle swarm optimization (PSO). Tajiki et al. [9] proposed a computationally efficient congestion avoidance scheme, called CECT, for software-defined cloud data centers. Tajiki et al. [10] focus on the problems of traffic engineering, failure recovery, fault prevention, and SFC with reliability and energy consumption constraints in Software Defined Networks (SDN).

2.2 OOD Detection

OOD detection is an emerging research direction. It can be achieved by various methods and has broad application prospects. In 2018, Dan Hendricks et al. [11] first proposed an OOD detection baseline based on deep learning and conducted many follow-up studies. The sub-health state of heavy machinery is a problem that has been ignored. Cui et al. [12] designed an OOD detection model with auxiliary modules to solve this problem. Wellhausen et al. [13] proposed to overcome overconfidence by using anomaly detection on multi-modal images for traversability classification, which is easily scalable by training in a self-supervised fashion from robot experience. Zhou et al. [14] used a contrast loss, which can improve the compactness of the representation so that OOD instances can be better distinguished from instances in the distribution. Auto-encoder is a model often used in OOD detection, but it has poor generalization ability to noise data. Zhang et al. [15] designed two auto-encoders for better noise immunity.

2.3 pros/cons table

We described the HDM-related work, and compiled them in **Table 1** to reflect the variations between them.

Table 1. The apparent differences of HDM compared to each category of the literature

Related work	Pros	Cons
[7]	Able to deal with nonlinear problems	Poor voice recognition ability
[8]	Flexibly decide the hyperparameters	Cannot be applied to devices with high stability
[9]	Minimizes network congestion	—
[10]	Applicable to large-scale networks	—
[11]	Founder of OOD detection	Performance is too weak

[12]	Higher fault early warning capability	Complex structure
[13]	Able to handle multimodal problems	It can only be used for anomaly detection of images
[14]	Overcome the semantic transfer problem	Occupy a lot of system resources
[15]	Have generalization ability	Poor classification accuracy

3. Model structure

3.1 Network Architecture

The proposed model is shown in Fig. 1. The model is composed of two VAEs with similar structures (parameters are shown in Table 2) [16], which train A (normal data) and B (fault data) in the input data, respectively. In the training process, first, map A to low latitude space, then determine the normal distribution \hat{X}_A through continuous training, store the extracted features in the register, and finally train B data. When the B data is mapped to the low dimensional space, the features in the register are used for filtering, and similar ones are filtered out, which eliminates the interference of normal samples and highlights the fault features. After training, the normal distribution \hat{X}_B can be obtained. During the test, the data output of the decoder is compared with the original data, and the reconstruction probability is obtained through (1). Finally, the result is obtained by the discriminator.

$$R = 1 - \frac{\sum x^2}{\sum A^2} \times 100\%, \quad x = A - A' \quad (1)$$

The R is the reconstruction probability, the A is the original data, and the A' is the reconstructed data.

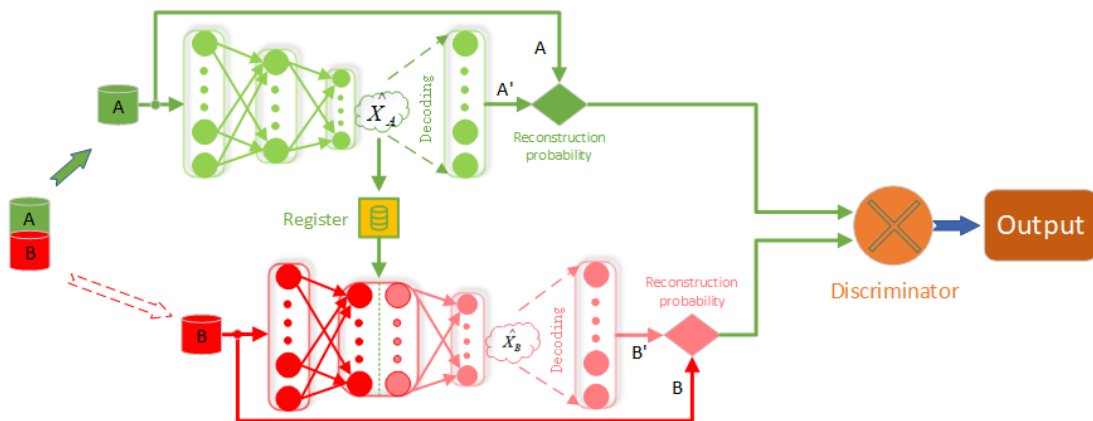


Fig. 1. The HDM structure diagram. A and B represent the two categories of data in the training set, respectively

Table 2. The parameters of the VAE generator.

Serial number	Types	Number of neurons
1	Input	1024
2	Fully connected layer	256
3	Fully connected layer	51
	Fully connected layer	5
4	Sampled latent variables	5
5	Fully connected layer	256
6	Output	1024

3.2 Algorithm and analysis

The two VAEs that make up the proposed model have the ability of independent classification, which use the reconstruction error to classify (as shown in [algorithm 1](#)). The mutual exclusion algorithm is used to determine the final classification of the two reconstruction errors.

Algorithm 1 Variational autoencoder classification algorithm

INPUT : Training dataset A , Test dataset $x^{(i)}$ $i = 1, \dots, N$, threshold a

OUTPUT: reconstruction probability $p_{\theta}(x|\hat{x})$

$\emptyset, \theta \leftarrow$ train a variation autoencoder using the training dataset X

for $i = 1$ **to** N **do**

$\mu_z(i), \sigma_z(i) = f_{\theta}(z|x^{(i)})$

draw L samples from $z \sim N(\mu_z(i), \sigma_z(i))$

for $l = 1$ **to** L **do**

$\mu_{\hat{x}}(i, l), \sigma_{\hat{x}}(i, l) = g_{\emptyset}(x|z^{(i, l)})$

end for

$reconstruction\ probadility(i) = \frac{1}{L} \sum_{l=1}^L p_{\theta}(x^{(i)} | \mu_{\hat{x}}(i, l), \sigma_{\hat{x}}(i, l))$

if $reconstruction\ probadility(i) < a$ **then**

$x^{(i)}$ is category A

else

$x^{(i)}$ is category B

end if

end for

In the process of predicting events by the model, assuming that A and B are mutually exclusive events [1], and the probability of occurrence of the event is P, then:

$$\begin{aligned} P(A+B) &= P(A)+P(B) \\ P(A)+P(B) &= 1 \end{aligned}$$

The more types of features the convolutional neural network model learns, the worse the effect [17], so we use VAE to train A and B separately. In the case of learning only one feature, the best recognition results for A and B are achieved, respectively.

Assuming that the model is M , the models M_A M_B are obtained after pre-training using A and B, respectively. The reconstruction probability of identifying A and B is respectively:

$$\begin{aligned} M_A(A) &= a \\ M_B(B) &= b \\ \because A \text{ and } B \text{ are mutually exclusive} \\ \therefore A &= \bar{B} \\ \therefore A &\rightarrow M_B \\ = \bar{B} &\rightarrow M_B = 1 - (M_B(B)) = 1 - b \end{aligned}$$

Hypothesis: the input data is C, and the weighting coefficients are 50%. Using the characteristics of mutually exclusive events that are not A or B, then the reconstruction probability $M(C)$ of the HDM is:

$$M(C) = \begin{cases} A \Rightarrow \frac{a+(1-b)}{2}, M_A(C) \geq M_B(C) \\ B \Rightarrow \frac{b+(1-a)}{2}, M_A(C) < M_B(C) \end{cases} \quad (2)$$

We infer that in the binary classification recognition of neural networks, two different classifications can also be regarded as mutually exclusive events. The training data used for binary classification is rarely mutually exclusive events. Still, in machine cognition, the essence of binary classification is to input two different data types and then output one of the two categories. As an independent event, the binary classification model has only two possibilities. They can be regarded as mutually exclusive events in a specific range, and the use of HDM will also increase the recognition effect. The example diagram is shown in Fig. 2.

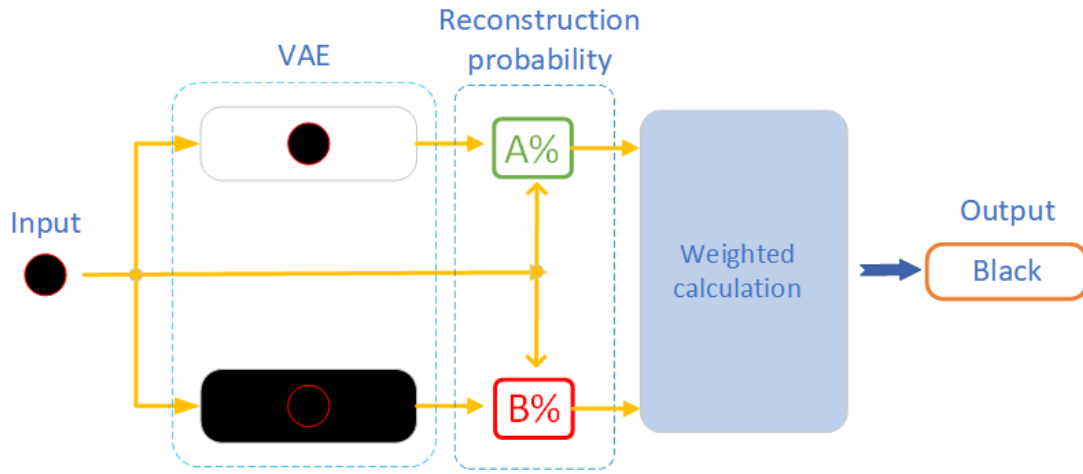


Fig. 2. The VAE model is trained to recognize black and white separately and create a schematic flow diagram when the input is black.

3.3 Contrast Model

Using the same data and parameters, the performance advantage of the double model based on mutual exclusion over the homologous single model will be demonstrated by comparing the results of HDM and VAE. In addition to proving the effectiveness of the principle we proposed, we must also demonstrate its progressiveness, so five widely used models are used for comparison.

3.3.1 ResNet

ResNet uses data preprocessing and a BN (batch normalization) layer in the network to address the gradient disappearance or explosion issue. In order to solve the degradation problem in the deep network, some layers of ResNet skip adjacent layer neurons to connect with the next layer, weakening the strong connection between each layer. The ResNet [18] ranks first in the DCASE2020 TASK2 single mode. The ResNet structure is shown in Table 3.

Table 3. ResNet structure parameter diagram, where x and y are used to control the receptive field of the network.

Type	Kernel number	kernel size 1	kernel size 2
Conv	2 ⁰	5	
BN	-	-	
Residual Block	2 ⁰	3	1
Max Pool	-	2	-
Residual Block	2 ⁰	3	3
Max Pool	-	2	-
Residual Block	2 ⁰	3	x
Residual Block	2 ⁰	3	y
Max Pool	-	2	-
Residual Block	2 ¹	1	1
Residual Block	2 ²	1	1

Residual Block	2 ²	1	1
Conv	1	1	-
Batch Normalization	-	-	-
Global Average Pooling	-	-	-

3.3.2 CAE

AE is a special neural network architecture, and the input and output are the same architecture. It is trained in an unsupervised method to obtain the lower dimensional expression of the input data. These low-latitude information expressions are reconstructed back to high-dimensional data expressions. CAE [19] replaces the Hessian matrix of AE with the Jacobian matrix, and other parts are almost the same. The mathematical expression of the Jacobian matrix is calculated in (3)

$$\|J_f(x)\|_F^2 = \sum_{i=1}^{d_h} (h_i(1-h_i))^2 \sum_{j=1}^{d_x} W_{ij}^2 \quad (3)$$

3.3.3 MobileFaceNet

The Google team has put forth MobilenetV2, which focuses on the compact CNN network in mobile or embedded devices. When the accuracy rate is only slightly decreased, the parameters and amount of calculation are considerably lowered. MobileFaceNet [20] has made five improvements based on MobileNetV2: separable convolution instead of average pool layer, uses Insightface loss function for training, reduces channel expansion multiple, Prelu instead of relu, and employs batch normalization. The structure is shown in Table 4.

Table 4. MobileFaceNet network structure diagram.

Operator	Multiple	Channel	Repeat	Stride
Conv 2D	-	16	1	2
Bottleneck	1	8	1	1
	6	16	2	2
	6	16	3	2
	6	32	4	2
	6	48	3	1
	6	80	3	2
	6	160	1	1
Conv 2D	-	1280	1	1
Avg Pool	-	1280	1	-
Dense	-	num classes	1	-

3.3.4 WaveNet

WaveNet [21] model is a sequence generation model, which can directly learn the mapping of sampling value sequence, so it has a good synthesis effect. At present, WaveNet is applied in speech synthesis, acoustic model modeling, and vocoder and has excellent potential in speech synthesis. The structure is shown in Fig. 3.

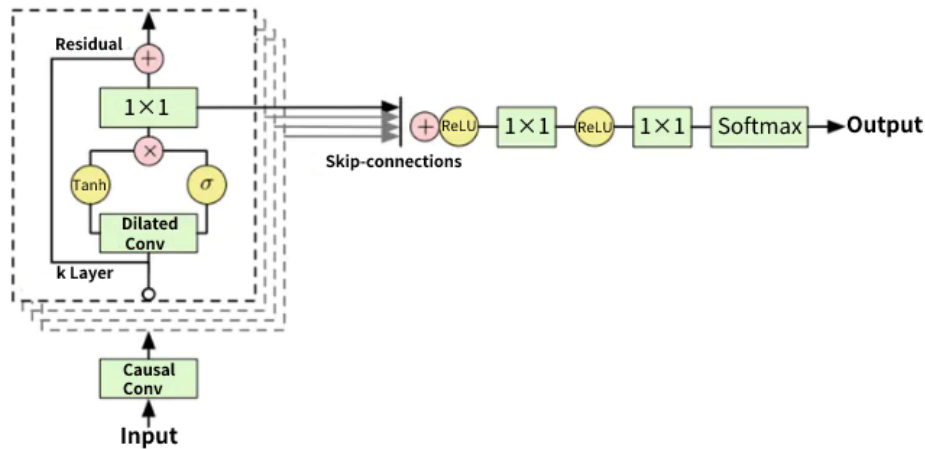


Fig. 3. Overview of the WaveNet entire architecture.

3.3.5 DCASE2020 TASK2 Baseline

The baseline system is a demonstration model provided by the DCASE2020 organizer. The reconstruction error of AE is compared with the threshold. It is abnormal if it is greater than the threshold, and normal if it is less than the threshold. The AE's hyper-parameters were as follows: epochs: 100, batch size: 512, Optimizer: Adam, learning rate: 0.001. The structure is shown in [Table 5](#).

Table 5. The parameters of the DCASE2020 TASK2 Baseline.

Serial number	Types	Number of neurons
1	Input	640
2	Dense layer	128
3	Dense layer	128
4	Dense layer	128
5	Dense layer	128
6	Bottleneck layer	8
7	Dense layer	128
8	Dense layer	128
9	Dense layer	128
10	Dense layer	128
11	Output	640

4. Experimental Results and Analysis

4.1 Experimental Data

This section introduces the data set, the evaluation criteria used in the experiment, and the models used for comparative experiments. DCASE2020 dataset was used to detect the performance of the proposed model. Datasets are ToyADMOS and MIMII, both of which were single-channel recordings. The down-sampling rate of all audio clips was 16kHz, and the length was about 10s. The normal sound sample data used in the TASK2 are divided into six categories: Toy-Car, Toy, Valve, Pump, Fan, and Slider. The first two are from Toy machines, while the rest are from real machines.

4.2 Experimental Evaluation Index

The OOD test has two fixed test indicators: true positive rate (TPR) and false-positive rate (FPR). They are calculated in (4) and (5), where TP and FN represent true positives and false negatives, and FP and TN represent false positives and true negatives.

$$TPR = \frac{TP}{TP + FN} \quad (4)$$

$$FPR = \frac{FP}{FP + TN} \quad (5)$$

To evaluate the advantages and disadvantages of the mechanical fault diagnosis model, the first criterion is to distinguish the performance of positive samples and negative samples. The area under the curve (AUC) value is between 0 and 1. The larger the value, the better the performance. The second criterion is the false positive rate. The partial-AUC (PUC) [22] value is calculated according to a part of the ROC [23] curve within a predetermined range.

4.3 Data preprocessing

It is important to note that the data must be preprocessed in accordance with [24] before being entered into the VAE model. In audio processing, the log-Mel filter is used for feature extraction. The filter can be used to obtain the energy distribution, which can then be used as an output feature. This feature is unaffected by audio properties and has a good recognition effect when the signal-to-noise ratio is low. Fig. 4 shows the log-Mel spectrogram of each material we enumerated.

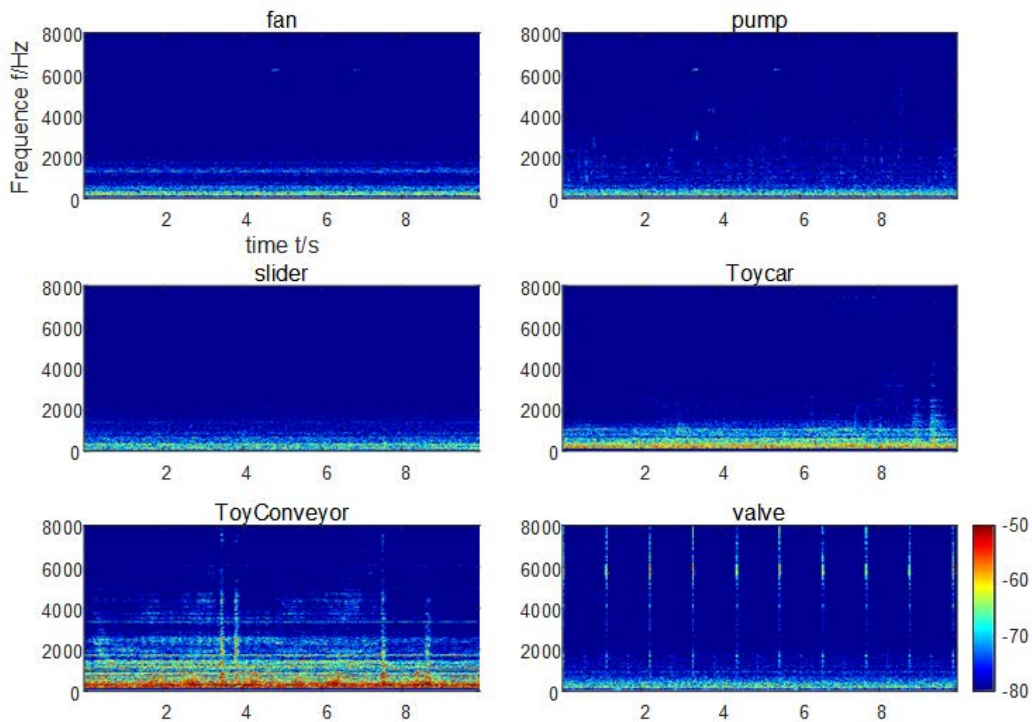


Fig. 4. Various data Log-Mel spectrogram. The horizontal axis represents time, and the vertical axis represents the frequency

4.4 Performance Evaluation

4.4.1 Characteristic filtering performance evaluation

In order to show the effect of the fault feature filtering module, the fault data of a ToyConveyor is shown as an example (as shown in Fig. 5). The comparison chart intuitively indicates that the filtering module has a filtering effect, weakening the feature points of normal data, making the fault characteristics more obvious.

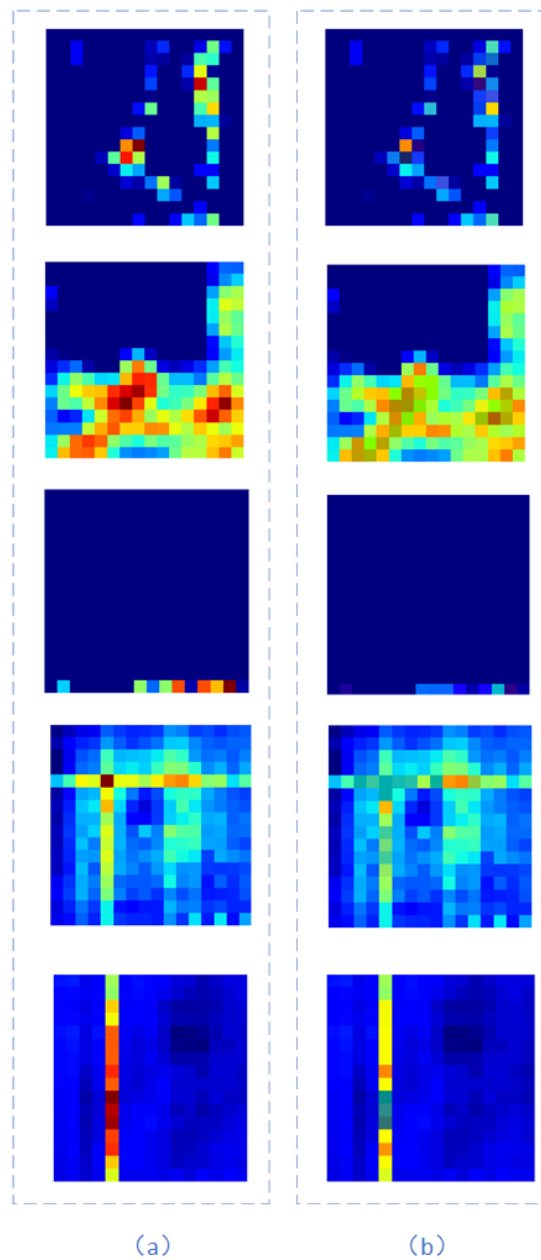


Fig. 5. Partial fault feature filtering effect. (a) is the original feature, (b) is the filtered feature.

4.4.2 Network performance evaluation

All experiments in this paper were conducted under the following setup: Intel (R) I9-10900X 3.70 GHz \times 10 CPU, NVIDIA RTX 3090 GPU with 24G memory \times 2, and PyTorch framework. According to previous experiment experience, the neural network's hyper-parameters were as follows: the initial learning rate was set to 0.00001, the momentum was set to 0.9, and the Adam optimizer was adopted. λ in the loss function was set to 0.0001.

In this section, the performances of the HDM and VAE are compared, and the structure of VAE is the same as that in [Table 1](#). As shown in [Table 6](#) and [Fig. 6](#), the AUC value and pAUC value of HDM exceed VAE in all projects. The fan project's AUC and pAUC of HDM are 14.94% and 26.57% higher than VAE, respectively. In the valve project, HDM has the least effect on VAE, and the AUC value is only increased by 0.33%. HDM's average AUC and pAUC are 6.07% and 10.71% higher than VAE, respectively.

Table 6. Comparison between HDM and homologous single model VAE.

Model		Fan	Pump	Slider	Valve	ToyCar	ToyConveyor
HDM	AUC	95.88%	96.65%	97.50%	91.37%	86.25%	88.47%
	pAUC	93.12%	91.14%	82.45%	89.63%	82.61%	71.39%
VAE	AUC	80.94%	85.26%	95.64%	91.04%	83.39%	83.46%
	pAUC	66.55%	74.35%	80.74%	77.10%	81.46%	68.98%

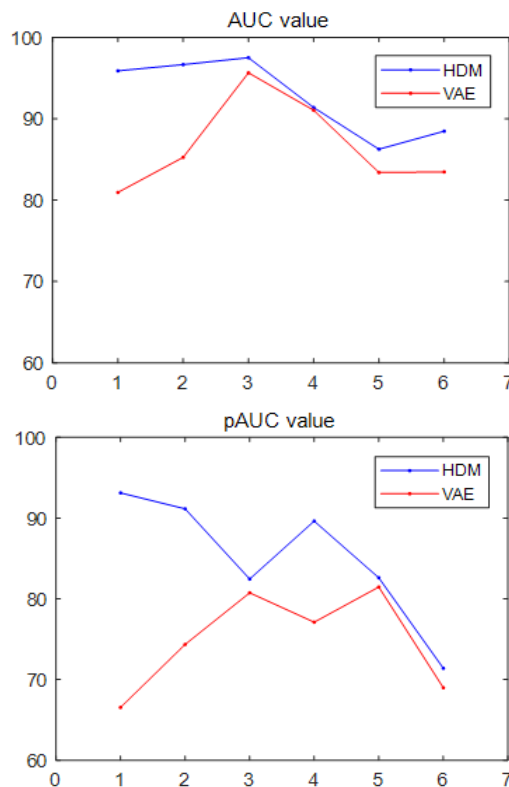


Fig. 6. Comparison between HDM and single homologous model. The vertical axis represents the size of the value, and the horizontal axis from 1 to 6 represents Fan, Pump, Slider, Valve ToyCar, and ToyConveyor in turn.

Table 7 present the AUC and pAUC results of different methods. The results show that in the six types of fault diagnosis of the DCASE2020 TASK2 task, the recognition effect of the HDM is better than the other five models. As shown in **Fig. 7**, the average AUC of the HDM is 14.78% higher than that of baseline, and ResNet has the best performance except for HDM, which is only 1.17% lower than the average AUC of the HDM.

Table 7. Comparison of the effects of HDM in the detection of real mutually exclusive events. All values are in %

AUC/pAUC	Baseline	WaveNet	CAE	MobileFaceNet	ResNet	HDM
Fan	82.80/65.80	93.51/85.10	81.82/76.98	94.58/92.84	93.65/82.47	95.88/93.12
Pump	82.37/64.11	95.87/89.53	88.17/80.36	90.05/76.55	93.68/91.20	96.65/91.14
Slider	79.41/58.87	97.36/94.60	86.49/74.65	93.56/88.51	97.29/88.74	97.50/82.45
Valve	57.37/50.79	87.94/91.58	84.59/62.41	73.77/66.26	90.15/86.65	91.37/89.63
ToyCar	80.14/66.17	86.73/89.30	91.25/87.36	85.92/80.59	86.06/70.01	86.25/82.61
ToyConveyor	85.36/66.96	87.22/72.59	72.23/60.23	86.17/70.88	88.28/79.15	88.47/71.39
average value	77.91/62.12	91.44/87.12	84.10/73.67	84.34/79.27	91.52/83.37	92.69/85.07

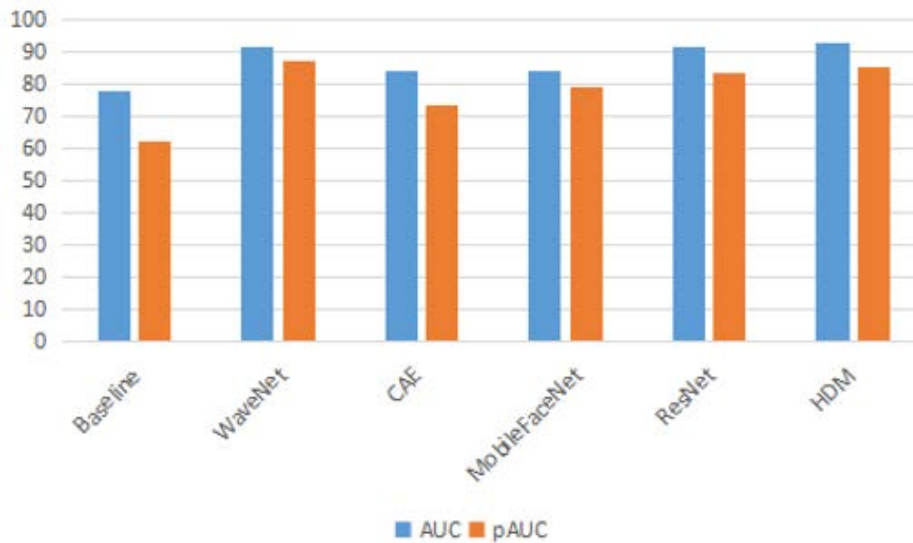


Fig. 7. Comparison of mean AUC and pAUC in six models.

The performance improvement is mainly attributed to three aspects: VAE has good learning ability for a single type of feature, the enhanced effect of mutual exclusion theory mutual verification, and the filtering mechanism of fault data characteristics are added. These compelling results verify that the HDM is better than the traditional network in binary classification and has good performance in mechanical fault diagnosis.

5. Conclusion

In this work, we propose the HDM method, which can effectively improve the performance of a single model for machine fault diagnosis. We take advantage of the high-fitting,

low-interference characteristics of single-type features and mutual verification of mutually exclusive events, integrating the principle of OOD detection in the output layer. The fault data feature filtering module is added, making the feature more prominent and easier to distinguish. Our method has achieved good recognition results in machine fault diagnosis data sets. The HDM method is currently only applicable to binary classification problems. Binary classification accounts for a minor portion of classification problems. In the future, we intend to decompose multi classification problems into multiple binary classification problems based on type similarity, and then to make HDM applicable to all classification problems.

References

- [1] N. Fenton, M. Neil, D. Lagnado, W. Marsh, B. Yet, A. Constantinou, "How to model mutually exclusive events based on independent causal pathways in Bayesian network models," *Knowledge-Based Syst.*, vol. 113, pp. 39-50, 2016. [Article \(CrossRef Link\)](#)
- [2] A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84-90, 2017. [Article \(CrossRef Link\)](#)
- [3] Simonyan K, Zisserman A., "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556*, 2014. [Article \(CrossRef Link\)](#)
- [4] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *Proc. of IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 770-778, 2016. [Article \(CrossRef Link\)](#)
- [5] Logutov OG, Robinson AR., "Multi-model fusion and error parameter estimation," *Q J Roy Meteorol Soc.*, 131, 3397-3408. 2005. [Article \(CrossRef Link\)](#)
- [6] Liang, S., Li, Y., Srikant, R., "Enhancing The Reliability of Out-of-distribution Image Detection in Neural Networks," *Machine Learning*, arXiv:1706.02690, 2017. [Article \(CrossRef Link\)](#)
- [7] Tang, S., Yuan, S., Zhu, Y., "Convolutional neural network in intelligent fault diagnosis toward rotatory machinery," *IEEE Access*, 8, 86510–86519, 2020. [Article \(CrossRef Link\)](#)
- [8] Haidong, S., Ziyang, D., Junsheng, C., Hongkai, J., "Intelligent fault diagnosis among different rotating machines using novel stacked transfer auto-encoder optimized by PSO," *ISA Trans.*, 105, 308–319, 2020. [Article \(CrossRef Link\)](#)
- [9] Tajiki, M.M., Akbari, B., Shojafar, M. et al., "CECT: computationally efficient congestion-avoidance and traffic engineering in software-defined cloud data centers," *Cluster Comput.*, 21, 1881–1897, 2018. [Article \(CrossRef Link\)](#)
- [10] Tajiki M M, Shojafar M, Akbari B, et al., "Joint Failure Recovery, Fault Prevention, and Energy-efficient Resource Management for Real-time SFC in Fog-supported SDN," *arXiv:1807.00324*, 2018. [Article \(CrossRef Link\)](#)
- [11] Hendrycks, D., Gimple, K., "A Baseline for Detecting Misclassified and Out-of-Distribution Examples in Neural Networks," *arXiv:1610.02136*, 2017. [Article \(CrossRef Link\)](#)
- [12] Cui, P., Wang, JJ., Li, XB., Li, CF., "Sub-Health Identification of Reciprocating Machinery Based on Sound Feature and OOD Detection," *MACHINES*, 2021, 9(8), 179, 2021. [Article \(CrossRef Link\)](#)
- [13] L. Wellhausen, R. Ranftl and M. Hutter, "Safe robot navigation via multi-modal anomaly detection," *IEEE Robot. Automat. Lett.*, vol. 5, no. 2, pp. 1326-1333, Apr. 2020. [Article \(CrossRef Link\)](#)
- [14] Zhou, W., Chen, M., "Contrastive Out-of-Distribution Detection for Pretrained Transformers," *arXiv:2104.08812*, 2021. [Article \(CrossRef Link\)](#)
- [15] Zhang, J., Zhang, Y., Bai, L., Han, J., "Lossless-constraint denoising based auto-encoders," *Signal Process., Image Commun.*, vol. 63, pp. 92-99, 2018. [Article \(CrossRef Link\)](#)
- [16] Kingma, D. P., M. Welling, "Auto-Encoding Variational Bayes," *arXiv:1312.6114*, 2013. [Article \(CrossRef Link\)](#)
- [17] Buda M., Maki A., Mazurowski, M.A., "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Netw.*, vol. 106, pp. 249-259, 2018. [Article \(CrossRef Link\)](#)

- [18] He, K., Zhang, X., Ren, S., Sun, J., “Deep residual learning for image recognition,” in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016. [Article \(CrossRef Link\)](#)
- [19] Rifai, S., Vincent, P., Muller, X., Glorot, X., Bengio, Y., “Contractive auto-encoders: Explicit invariance during feature extraction,” in *Proc. of the 28th International Conference on Machine Learning (ICML 2011)*, pp. 833–840, 2011.
- [20] Sheng C., Yang L., Xiang G., “MobileFaceNets: Efficient CNNs for Accurate Real-time Face Verification on Mobile Devices,” in *Proc. of CCBR 2018: Biometric Recognition*, pp. 428-438, 2018. [Article \(CrossRef Link\)](#)
- [21] Oord A., Dieleman S., Zen H., “WaveNet: A Generative Model for Raw Audio,” *arXiv:1609.03499*, 2016. [Article \(CrossRef Link\)](#)
- [22] Bradley, P., “The use of the area under the ROC curve in the evaluation of machine learning algorithms,” *Pattern Recognition*, vol. 30, no. 7, pp. 1145-1159, 1997. [Article \(CrossRef Link\)](#)
- [23] T. Fawcett, “An introduction to roc analysis,” *Pattern Recogn. Lett.*, vol. 27, no. 8, pp. 861-874, 2006. [Article \(CrossRef Link\)](#)
- [24] Noda, K., Yamaguchi, Y., Nakadai, K., Okuno, H.G., Ogata, T., “Audio-visual speech recognition using deep learning,” *Appl. Intell.*, 42(4), pp. 722–737, 2015. [Article \(CrossRef Link\)](#)



PENG CUI He is a doctoral student at Yanshan University in China. His research interests include deep learning, intelligent fault diagnosis and sound event detection.



XUAN LUO She is an engineer in Hebei Petroleum University of Technology. Her current research interests include deep learning framework, network security and audio signal processing.



JING LIU She is currently a Ph.D. student in electronic science and technology at Yanshan University. Her main research interest is biological signal processing and synchronization analysis.